Plant Pathology (2019) **68**, 278–287 Doi: 10.1111/ppa.12946

Multilocus sequence typing analysis of Italian *Xanthomonas campestris* pv. *campestris* strains suggests the evolution of local endemic populations of the pathogen and does not correlate with race distribution

P. Bella[a]*, C. Moretti[b], G. Licciardello[c], C. P. Strano[c], A. Pulvirenti[d], S. Alaimo[d], M. Zaccardelli[e], F. Branca[c], R. Buonaurio[b], J. G. Vicente[f] and V. Catara[c]

[a]*Department of Agricultural, Food and Forest Sciences, University of Palermo, Viale delle Scienze, Ed. 4, 90128 Palermo;* [b]*Department of Agricultural, Food and Environmental Sciences - University of Perugia, Borgo XX Giugno 74, 06121 Perugia;* [c]*Department of Agriculture, Food and Environment, University of Catania, Via Santa Sofia, 100 - 95123;* [d]*Department of Clinical and Experimental Medicine, University of Catania, Viale Andrea Doria, 6 – 95125, Catania;* [e]*CREA - Horticulture and Floriculture Research Centre, Via dei Cavalleggeri, 25 - 84098, Pontecagnano (SA), Italy; and* [f]*School of Life Sciences, University of Warwick, Wellesbourne, CV35 9EF, UK*

Xanthomonas campestris pv. *campestris* (Xcc) is the causal agent of black rot in Brassicaceae. It is widespread in Italy and severe outbreaks occur under conditions that favour disease development. In this study a multilocus sequence typing approach (MLST) based on the partial sequence of seven loci was applied to a selection of strains representative of the main areas of cultivation and hosts. The aim was to investigate whether the long tradition of brassica crops in Italy has influenced the evolution of different Xcc populations. All loci were polymorphic; 14 allelic profiles were identified of which 13 were unique to Italian strains. Based on the seven loci, the most common genotype within the Italian Xcc strains (AP1) was also the most representative genotype found in worldwide Xcc strains. This genotype was included in a new clonal complex in addition to three other clonal complexes already identified in Xcc populations. The phylogenetic reconstruction using a concatenated dataset of four conserved protein-coding genes, *dnaK*, *fuyA*, *gyrB* and *rpoD*, showed that the Italian strains belonged to two genetic groups. Physiological races were also investigated for the first time in Italy. The race structure of Xcc was determined by inoculating eight differential *Brassica* lines belonging to five species and showed that, in Italy, race 4 is the most widespread, followed by races 1 and 6. No correlation was found between allelic profiles, host of isolation, geographical origin and races, although a prevalent race was identified within the same clonal complex.

Keywords: black rot of Brassicaceae, MLST, population diversity, race designation, *Xanthomonas campestris* pv. *campestris*

Introduction

Xanthomonas campestris pv. *campestris* (Xcc) is the causal agent of black rot, a severe, seedborne, systemic, vascular disease of vegetable brassica crops (Williams, 1980; Vicente & Holub, 2013; CABI, 2017). Xcc is present worldwide and induces disease in a large number of Brassicaceae species, including economically important brassica crops and a number of other cruciferous crops, ornamentals and weeds. Typical disease symptoms include V-shaped yellow to dark-brown lesions starting from the margin of the lower leaves and blackening of the veins that result from bacterial colonization of the xylem vessels. The affected tissues can become necrotic, with leaves falling prematurely; systemic infections can lead to stunted growth and the death of young plants. Abortion of seeds may occur when pods are infected during the early stage of their development. Otherwise, the pods may show dark and shrivelled irregular areas.

In addition to the pathovar *campestris*, another five pathovars have been described within the *Xanthomonas campestris* infecting Brassicaceae: *aberrans*, *armoraciae*, *barbareae*, *incanae* and *raphani* (Vicente & Holub, 2013). On the basis of pathogenicity tests and multilocus sequence analysis (MLSA) it was proposed that only three pathovars induce three different diseases on *Brassica* plants: *X. campestris* pv. *campestris*, causal agent of black rot of crucifer crops, *X. campestris* pv. *raphani*, causal agent of leaf spot disease on Brassicaceae and Solanaceae and *X. campestris* pv. *incanae*, causal agent of bacterial blight on ornamental plants (Fargier &

*E-mail: patrizia.bella@unipa.it

Published online 2 November 2018

Plant Pathology

Manceau, 2007; Fargier *et al.*, 2011). Multilocus sequence typing analysis (MLST), based on the partial sequence of eight conserved protein-coding genes, on a worldwide collection of 42 strains within the pathovars of *X. campestris* detected 23 sequence types (STs), with the six pathovars of *X. campestris* showing different allelic profiles (Fargier *et al.*, 2011). Phylogenetic evolution based on four genes (*dnaK*, *fuyA*, *gyrB* and *rpoD*) showed that *X. campestris* forms a monophyletic group that is split into two clusters, including strains previously identified as pathovars *armoraciae* and *aberrans* (Fargier & Manceau, 2007; Fargier *et al.*, 2011).

Nine physiological races have been identified within Xcc based on the response of different bacterial isolates on eight differential *Brassica* lines (Vicente *et al.*, 2001; Fargier & Manceau, 2007; Vicente & Holub, 2013). A simple gene-for-gene model based on the interaction of at least five avirulence genes and the matching *R* genes has been used to explain the relationship between Xcc races and the *Brassica* differentials (Vicente *et al.*, 2001; Fargier & Manceau, 2007; Vicente & Holub, 2013). Race definition is important to identify sources of resistance, to establish breeding programmes and also to set up effective strategies to control black rot. Many efforts have aimed to identify the race structure of Xcc populations in the United Kingdom (Vicente *et al.*, 2001), Portugal (Vicente, 2004), Nepal (Jensen *et al.*, 2010), northwestern Spain (Lema *et al.*, 2012), East Africa (Mulema *et al.*, 2012), southern Mozambique (Bila *et al.*, 2013) and India (Rathaur *et al.*, 2015; Singh *et al.*, 2016). Races 1 and 4, which have been described worldwide, are the most common races (Vicente *et al.*, 2001; Jensen *et al.*, 2010; Lema *et al.*, 2012; Mulema *et al.*, 2012; Vicente & Holub, 2013; Singh *et al.*, 2016). Two possible new races, 10 and 11, were described by Cruz *et al.* (2017), which deserve further analysis.

Xcc is present and widespread in brassica cultivations in Italy (CABI, 2017), both in production fields and in ornamental plants (Caponero & Iacobellis, 1994; Scortichini *et al.*, 1994; Catara *et al.*, 1999; Buonaurio *et al.*, 2003). When the climatic conditions are conducive to the disease, severe black rot epidemics have been reported, leading to 100% disease incidence and conspicuous yield losses. PCR-based fingerprinting methods have highlighted a high variability in a collection of 141 Italian Xcc strains isolated from different brassica species, with 43 and 21 haplotypes identified by M13 and BOX primers, respectively (Zaccardelli *et al.*, 2008). Unfortunately, no information is available on the race composition of Italian populations of Xcc except for four strains assigned to race 4 within the framework of a worldwide study (Vicente *et al.*, 2001).

In Italy, the cultivation of brassica crops is mainly concentrated in the central and southern areas and in Sicily. Cauliflower and broccoli are the most represented species, for which Italy is ranked fifth worldwide (FAOSTAT, 2014; ISTAT, 2016). A notable diversity of cauliflower and broccoli-like vegetables has developed in Europe, which probably originated from highly localized crops in Italy; these, in turn, may have originated from germplasm introduced in Roman times from the eastern Mediterranean (Dixon, 2007). Local biodiversity is often maintained by the farmers themselves, as there are still many different landraces that are cultivated in both large fields and home gardens, which are savoured by local people due to their organoleptic traits (Branca, 2008; Ciancaleoni *et al.*, 2014). This biodiversity is very important for breeding programmes both in terms of the specific morphological and organoleptic traits as well as the resistance traits to biotic and abiotic stresses.

One of the main goals of this study was to investigate whether the long tradition of brassica crops in Italy has influenced the evolution of different Xcc populations. An MLST approach was used that is easily comparable between laboratories, as DNA sequence data can be easily made available. The MLST scheme applied in this study was previously validated to study the genetic relatedness of the six originally described *X. campestris* pathovars associated with Brassicaceae (Fargier *et al.*, 2011).

In order to analyse representative strains, samples were selected from the collection of 141 strains characterized by Zaccardelli *et al.* (2008) from the most important brassica-producing zones in central and southern Italy.

Unfortunately, no correlation to date has been reported between DNA fingerprinting, MLST allelic profiles or STs and Xcc races (Lange *et al.*, 2016). Thus, to provide useful information for breeding programmes or landrace resistance, the race structure in Italy was investigated according to Vicente & Holub (2013), using eight differential cultivars and accessions.

Materials and methods

Bacterial strains

The set of strains of Xcc used in this study is listed in Table 1. The strains, isolated from black rot-affected Brassicaceae plants, represent a selection of 21 strains from a collection of 141 strains from central and southern Italy already characterized by PCR-based DNA fingerprinting (Zaccardelli *et al.*, 2007, 2008); in addition, 10 strains were isolated in Sicily and are in the collection of the University of Catania. Overall, strains were collected from six regions representing more than 80% of the Italian production of brassicas and from eight different brassicas (broccoli, cabbage, cauliflower, seakale, kale, kohlrabi, Savoy cabbage and rutabaga) over a period of approximately 15 years (data not shown). The identification of these strains was confirmed by PCR with Xcc-specific primers based on the *hrcC* gene (Zaccardelli *et al.*, 2007; data not shown). Representative isolates of Xcc races 1, 4 and 6 obtained from the LMG and WHRI culture collections were included for comparative purposes (Table 1). All strains were routinely grown at 28 °C on nutrient agar (Oxoid) supplemented with 1% dextrose (NDA). For long-term storage, bacteria were maintained in nutrient broth with 15% glycerol at −80 °C.

Table 1 *Xanthomonas campestris* pv. *campestris* (Xcc) strains used in this study with their race identification and allelic profiles.

Strain	Host	Geographical origin	Race	Genotype	Allelic profile[a]							Clonal complex[b]
					atpd	*dnaK*	*efp*	*fyuA*	*glnA*	*gyrB*	*rpoD*	
ISPaVe 016	Kohlrabi	Lazio, Italy	1	AP8	4	1	6	13	10	1	4	CC1
ISPaVe 1032	Cabbage	Lazio, Italy	4	AP8	4	1	6	13	10	1	4	CC1
IPV-NA 38	Cauliflower	Lazio, Italy	4	AP8	4	1	6	13	10	1	4	CC1
DAPP-PG 366	Cauliflower	Apulia, Italy	1	AP4	7	1	4	4	10	1	4	CC2
DAPP-PG 355	Cauliflower	Marche, Italy	4	AP4	7	1	4	4	10	1	4	CC2
PVCT 189.1.1	Cabbage	Sicily, Italy	4	AP4	7	1	4	4	10	1	4	CC2
PVCT 190.1.1	Cauliflower	Sicily, Italy	4	AP4	7	1	4	4	10	1	4	CC2
PVCT 72.1	Kohlrabi	Sicily, Italy	4	AP4	7	1	4	4	10	1	4	CC2
PVCT 62.1	Kohlrabi	Sicily, Italy	4	AP5	7	1	N1	4	10	1	4	CC2
PVCT 62.4	Kohlrabi	Sicily, Italy	4	AP13	7	1	N1	4	10	1	N1	CC2
PVCT 63.1	Kohlrabi	Sicily, Italy	4	AP14	7	1	4	4	10	1	N2	CC2
ISCI 1	Kohlrabi	Campania, Italy	4	AP2	2	1	6	14	1	1	11	CC3
ISCI 42	Kale	Campania, Italy	4	AP2	2	1	6	14	1	1	11	CC3
ISCI 45	Savoy cabbage	Campania, Italy	4	AP2	2	1	6	14	1	1	11	CC3
ISCI 122	Kale	Campania, Italy	1	AP12	4	9	6	14	1	3	11	CC3
DAPP-PG 306	Cauliflower	Umbria, Italy	1	AP1[c]	11	1	1	14	1	1	11	CC4
DAPP-PG 308	Cauliflower	Marche, Italy	1	AP1	11	1	1	14	1	1	11	CC4
DAPP-PG 398	Rutabaga	Marche, Italy	1	AP1	11	1	1	14	1	1	11	CC4
ISCI 12	Cauliflower	Campania, Italy	1	AP1	11	1	1	14	1	1	11	CC4
ISCI 88	Cabbage	Campania, Italy	1	AP1	11	1	1	14	1	1	11	CC4
OMP-BO 588/90	Cabbage	Emilia Romagna, Italy	1	AP1	11	1	1	14	1	1	11	CC4
PVCT 75.1.1	Broccoli	Sicily, Italy	1	AP1	11	1	1	14	1	1	11	CC4
PVCT 166/96	Cauliflower	Sicily, Italy	6	AP1	11	1	1	14	1	1	11	CC4
IPV-NA 39	Kohlrabi	Apulia, Italy	4	AP6	1	1	1	14	1	1	12	CC4
PVCT 67.2	Kohlrabi	Sicily, Italy	4	AP6	1	1	1	14	1	1	12	CC4
PVCT 190.2.1	Kale	Sicily, Italy	6	AP7	11	1	1	14	1	1	12	CC4
DAPP-PG 249	Cabbage	Umbria, Italy	4	AP3	4	N1	4	4	N1	1	11	S
ISCI 127	Kale	Campania, Italy	4	AP3	4	N1	4	4	N1	1	11	S
DAPP-PG 330	Cauliflower	Umbria, Italy	6	AP9	2	N2	6	N1	1	1	12	S
ISCI 31	Broccoli	Campania, Italy	6	AP10	7	N3	1	N2	1	1	12	S
ISCI 70	Seakale	Campania, Italy	4	AP11	2	1	6	14	10	1	N1	S
Reference strains												
LMG 8001	Cabbage	United Kingdom	1	AP15	11	1	1	14	1	1	N3	CC4
WHRI 3811	*Brassica oleracea*	USA	1	nt	nt	1	nt	14	nt	1	11	nt
WHRI 1279A	Cabbage	United Kingdom	4	nt	nt	1	nt	4	nt	1	4	nt
WHRI 6181	*Brassica rapa*	Portugal	6	nt	nt	1	nt	N3	nt	1	12	nt

nt, not tested. DAPP-PG: Department of Agricultural, Food and Environmental Sciences - University of Perugia, Italy; ISCI: CREA - Horticulture and Floriculture Research Centre, Battipaglia, SA, Italy; IPV-NA: Department of Agricultural Sciences, University of Naples, Naples, Italy; ISPAVe: Research Centre for Plant Protection and Certification, Council for Agricultural Research and the Analysis of Agricultural Economics, Rome, Italy; OMP-BO: Plant Protection Service, Emilia Romagna, Bologna, Italy; PVCT: Department of Agriculture, Food and Environment, University of Catania, Italy; LMG: Culture Collection Laboratorium voor Microbiologie, University of Ghent, Belgium; WHRI: School of Life Sciences, University of Warwick, Wellesbourne, UK.

[a]Alleles identical to those identified by Fargier *et al.* (2011) were assigned the number given by those authors; new alleles detected in this study were assigned a number preceded by the letter 'N'.

[b]Clonal complexes (CC) and singletons (S) within 15 allelic profiles (AP) of Xcc strains analysed in this study and 13 sequence types (ST) of Xcc from Fargier *et al.* (2011) were defined by eBURST v. 3. Information on CCs of Xcc strains from the international collection is available in Table S4.

[c]Corresponding to ST17 of Fargier *et al.* (2011).

DNA isolation and sequencing

Genomic DNA was extracted from 1.5 mL bacterial cultures grown in nutrient broth on a rotatory shaker for 48 h at 27 °C according to Licciardello *et al.* (2011). DNA concentration was measured with a NanoDrop spectrophotometer. Seven genes, including the six housekeeping genes *atpD* (ATP synthase subunit beta), *dnaK* (heat shock protein 70, molecular chaperone DnaK), *efp* (elongation factor P), *glnA* (glutamine synthetase I), *gyrB* (DNA gyrase subunit B) and *rpoD* (RNA polymerase sigma-70 factor), as well as one gene coding for a transmembrane protein, *fyuA* (TonB-dependent receptor), were amplified according to Fargier *et al.* (2011) with minor modifications. All amplifications were carried out in a final volume of 40 μL containing 1× Colorless GoTaq Flexi buffer (Promega), 1.5 mM $MgCl_2$, 0.2 mM of each dNTP (Invitrogen), 400 nM of each primer, 1 U GoTaq Flexi DNA polymerase (Promega) and 1 μL of DNA (50–100 ng) as template. An initial denaturation at 94 °C for 180 s was followed by 35 cycles, each consisting of 50 s at 94 °C, 50 s at 63 °C for *dnaK*, *rpoD* and *fyuA*, 60 °C for *gyrB* or 56 °C for *atpD*, *efp* and *glnA*, 1 min at 72 °C, and a final extension of 7 min at 72 °C. Amplification was confirmed on 1% agarose gels stained with GelRed (Biotium). PCR amplicons were sequenced on both

strands by BMR Genomics (Padova, Italy) using primers defined by Fargier *et al.* (2011).

Sequence analysis and phylogenetic studies

The consensus sequences for each sequenced gene were determined by assembling forward and reverse sequences using BioEdit v. 7.0.9 software and ambiguous sequences at both ends were removed. These sequences were aligned using MEGA v. 6.0 (Tamura *et al.*, 2013) and trimmed according to nucleotide length of sequences retrieved from GenBank of a worldwide collection of 42 *X. campestris* strains, causal agents of different diseases on Brassicaceae (Fargier *et al.*, 2011). The alleles at the seven loci (*atpD*, *dnaK*, *efp*, *fyuA*, *glnA*, *gyrB*, *rpoD*) were identified by comparing the sequences of Xcc strains with those of the alleles obtained by Fargier *et al.* (2011) at the same loci (Table S1). Each gene sequence was defined as a new allele if it differed at a single nucleotide site and was assigned to an allele number. The allelic profile (AP) was defined by combining the allele numbers for each isolate. BioNumerics v. 7.6 software (Applied-Maths) was used to define the different alleles within the Xcc strains used in this study and the corresponding allele profile and to build a minimum spanning tree. Groups of isolates with closely related APs were grouped within clonal complexes (CCs) using eBURST v. 3 (Feil *et al.*, 2004; http://eburst.mlst.net/). Xcc strains were assigned to the same group if they shared identical alleles at six out the seven loci with at least one other Xcc strain of the group.

Phylogenetic analyses were performed on the concatenated dataset of four selected genes *(dnaK*, *fyuA*, *gyrB*, *rpoD)* because no incongruent phylogenies were found between individual gene sequences and the concatenated dataset for these genes (Fargier *et al.*, 2011). Thirty-one Italian Xcc strains, four reference strains (LMG 8001, WHRI 3811, WHRI 1279A WHRI 6181) and a selection of sequences from 21 *X. campestris* pathovars from an international collection (Fargier *et al.*, 2011), chosen on the basis of difference in the allelic profile at seven genes, were included for phylogenetic analysis (Table S2). Neighbour-joining (NJ) trees were generated with MEGA v. 6.0 using the Jukes–Cantor distances method and 1000 bootstrap replicates. *Xanthomonas euvesicatoria* 85.10, *X. citri* subsp. *citri* 306, and *X. oryzae* pv. *oryzae* KACC10331 were included as out-groups. Total recombination and mutation frequencies were performed by SplitsTree v. 4.14 (Huson, 1998; Huson & Bryant, 2006). The split decomposition method was performed based on the concatenated sequences of all of four genes.

Race determination

The race of each strain was determined by inoculating eight differential *Brassica* lines belonging to five species (Vicente & Holub, 2013). Xcc strains were grown on NAD for 48 h at 28 °C. Cultures were suspended in sterile distilled water and the resulting bacterial suspensions were adjusted to 10^8 CFU mL^{-1}. Plants were sown in 24-cell foam trays and maintained in greenhouse with a temperature of 20–25/15–20 °C (day/night). Inoculations were performed according to Vicente *et al.* (2001). Briefly, the plants were inoculated 4 weeks after sowing by clipping the three youngest leaves of each plant at the secondary veins near the margins. Mouse-tooth forceps with the teeth wrapped in cotton wool and dipped into the bacterial suspension were used to perform 8–10 points of inoculation per leaf. Three plants were inoculated per strain. The number of infected points per leaf and the presence of symptoms were recorded 7 and 15 days after inoculations. The presence of typical V-shaped lesions was recorded as a compatible interaction, (+, susceptibility) whereas the absence of symptoms or observation of hypersensitive response (HR) at the inoculation point were assessed as an incompatible interaction (−, resistance). Xcc isolates that were compatible with differential cultivars Wirosa F1, Miracle F1, SxD1 (*Brassica oleracea*), Just Right Hybrid Turnip, Seven Top Turnip (*B. rapa*), and COB60 (*B. napus*), designated as race 1; isolates compatible with cultivars Wirosa F_1, SxD1 and Miracle F_1 were designated as race 4; and isolates that were compatible on all eigth cultivars were designated as race 6.

Results

Allelic profile analysis

The sequences of seven loci, namely *atpD*, *dnaK*, *efp*, *fyuA* *glnA*, *gyrB*, *rpoD*, were obtained for both strands from 31 Italian strains and one reference strain (LMG 8001). For strains WHRI 3811, WHRI 1279A, WHRI 6181, only *dnaK*, *rpoD*, *fyuA* sequences were obtained but the *gyrB* sequences were already present in GenBank and were used only for phylogenetic analysis (Table S3). The 31 sequences of each of the seven loci for the Italian and reference strains were aligned with the sequences of 21 strains obtained from GenBank related to the worldwide study on *X. campestris* crucifer-attacking pathovars (Fargier *et al.*, 2011). No gap and no insertion in the sequences of any strains of *X. campestris* were detected. The size of the seven alignments ranged from 434 to 697 bp (Table 2). A total of 233 newly generated sequences were deposited in GenBank under the accession numbers MH520847 to MH520878 (*atpD* fragment), KY379972 to KY380003 (*dnaK* fragment), MH520879 to MH520910 (*efp* fragment), KY499060 to KY499091 (*fyuA* fragment), MH520911 to MH520942 (*glnA* fragment), MF133462 to MF133493 (*gyrB* fragment), MF494705 to MF133493 (*rpoD* fragment) (Table S3). The GenBank accession numbers for the partial sequences of the Xcc strains WHRI 3811, WHRI 1279A and WHRI 6181 were MG190365 to MG190367 for *dnaK*, MG029455 to MG029457 for *fyuA* and MG190368 to MG190370 for *rpoD* (Table S3). All loci were polymorphic, and the number of polymorphic nucleotide sites varied from 44 for the most polymorphic locus (*rpoD*) to three for the least polymorphic locus (*efp*), which represented 6.3% and 0.7% of the fragment length, respectively (Table 2). New alleles of each gene were assigned an allele number. To facilitate comparison with previous sequence data, alleles with identical sequences to those obtained by Fargier *et al.* (2011) were assigned the same number given by those authors; alleles with new sequences were allocated a new number. The number of alleles at each locus for Italian strains were: two for *gyrB*, four for *efp*, three for *glnA*, five for *atpD*, *dnaK* and *fyuA*, and six for *rpoD*. The allelic profile was defined by the combination of allele numbers for each strain. Fourteen different genotypes were identified amongst the 31 Italian Xcc strains, six of which (AP 1–4, 6 and 8) were common to two to eight strains whereas the others (AP 5, 7, 9–14) were unique to only one Xcc strain. A different allelic profile was obtained for Xcc strain LMG 8001 (AP 15).

Table 2 Number of alleles and polymorphic sites at the seven loci analysed in 31 Italian strains and one reference strain of *Xanthomonas campestris* pv. *campestris*.

Locus	Function	Fragment size (bp)	No. of alleles			No. of polymorphic sites	% of variable sites
			Fargier *et al.* (2011)	This study	New alleles		
atpD	ATP synthase β chain	512	14	5	0	12	2.3
dnaK	Heat shock protein 70	540	10	5	3	6	1.3
efP	Elongation factor P	434	7	4	1	3	0.7
fyuA	TonB-dependent receptor	669	15	5	2	14	2.0
glnA	Glutamine synthetase	582	11	3	1	6	1.0
gyrB	Gyrase subunit B	521	6	2	0	8	1.5
rpoD	RNA polymerase sigma-70 factor	697	13	6	3	44	6.3
Concatenated		3955	21	15	14	93	2.6

The most common allelic profile was AP1 (alleles 11-1-1-14-1-1-11 for *atpD*, *dnaK*, *efp*, *fyuA*, *glnA*, *gyrB* and *rpoD*, respectively), which was shared by 25% of Italian Xcc strains obtained from different *B. oleracea* varieties in five regions in Italy (Table 1). AP1 was the only genotype detected in this study that had previously been identified in other Xcc strains isolated in other countries (Fargier *et al.*, 2011; Lange *et al.*, 2016). The AP4 profile was shared by five Italian Xcc strains and was the second most common. Thirteen out of 14 allelic profiles detected in this study were unique to the Italian strains.

Fifteen allelic profiles of the Xcc analysed in this study and 13 STs of Xcc from Fargier *et al.* (2011) were grouped into four clonal complexes and eight singletons (Tables 1 & S4). Three out of the four clonal complexes coincided with those previously identified and were designated accordingly (Fargier *et al.*, 2011). The new clonal complex (CC4) included ST17 previously described as a singleton as well as AP1, AP6, AP7 and AP15. Based on the seven loci scheme, ST17 and AP1 were identical and the most widespread. The founding genotypes of the clonal complexes CC2 and CC4, corresponding to the AP/ST that differed from the largest number of other APs/STs at only one single locus, were APs of the Xcc strains isolated in Italy. For the Italian Xcc strains, clonal complexes CC1 and CC3 grouped strains from the same geographical origin (CC1, Xcc from Latium; CC3, Xcc from Campania), whereas CC2 and CC4 included strains isolated from different regions. Seven of the Xcc strains isolated in Sicily belonging to race 4 were included in CC2. No correlation was found between clonal complex and host of isolation, because in each clonal complex the Xcc strains were isolated from different *Brassica* species. A minimum spanning tree showed the genetic relationships among 27 Xcc allelic profiles and four clonal complexes (Fig. 1). Most (58%) of the Italian Xcc strains were linked by single locus variation.

Phylogenetic analysis

Phylogenetic analysis was performed on the concatenated sequences of four loci (*dnaK*, *rpo*, *fyuA* and *gyrB)* of 31 Italian strains (14 genotypes), four reference strains (LMG 8001, WHRI 3811, WHRI 1279A and WHRI 6181) and 21 representative strains of *X. campestris* pathovars associated with Brassicaceae previously obtained by Fargier *et al.* (2011) (Tables S1 & S2). *Xanthomonas euvesicatoria* 85.10, *X. citri* subsp. *citri* 306 and *X. oryzae* pv. *oryzae* KACC10331 were used as outgroup strains. The phylogenetic tree constructed by the neighbour-joining method showed that Xcc strains isolated in Italy were included in the *X. campestris* monophyletic group described by Fargier *et al.* (2011) (Fig. 2). All *X. campestris* pathovars were grouped in two main clusters (I and II) including the 14 Xcc Italian genotypes. Xcc strains were divided between cluster I and cluster IIA. Cluster I included allelic profiles AP1–3, AP6–7, AP9–10 and AP12. Most of the Italian strains (19 out 31) were in cluster I. Eighteen strains were isolated from six varieties of *B. oleracea* (cauliflower, broccoli, kale, kohlrabi, cabbage, Savoy cabbage) and one strain from rutabaga (*Brassica napobrassica*) in all the regions investigated (accounting for more than 80% of the production of brassicas in Italy). Three Xcc reference strains from international collections (LMG 8001, WRHI 6181 and WRHI 3818) and eight STs from Fargier *et al.* (2011) were also grouped in cluster I. These strains were of worldwide origin (UK, Portugal, France, Spain, Belgium, Tanzania, USA and Australia) and were isolated from *B. oleracea*, *B. rapa* and *Iberis* sp. from 1967 to 2000.

Cluster II included the remaining genotypes of *X. campestris* pathovars associated with Brassicaceae, subdivided into three subclusters according to Fargier *et al.* (2011). Subcluster IIA grouped allelic profiles AP4–5, AP8, AP11, AP13–14 shared by 12 Italian Xcc strains isolated from three *B. oleracea* types (cauliflower, cabbage and kohlrabi) and one strain from seakale (*Crambe* sp.). Strains belonging to other *X. campestris* pathovars from Fargier *et al.* (2011) were in the two subclusters IIb and IIc. For the Italian Xcc strains, no correlation was found between geographical origin, host of isolation, race and position in the phylogenetic tree.

Split decomposition analysis of the concatenated gene sequences of four loci (*dnaK*, *rpoD*, *fyuA* and *gyrB)* showed a network-like structure with multiple parallel paths (Fig. S1). The results are coherent with those

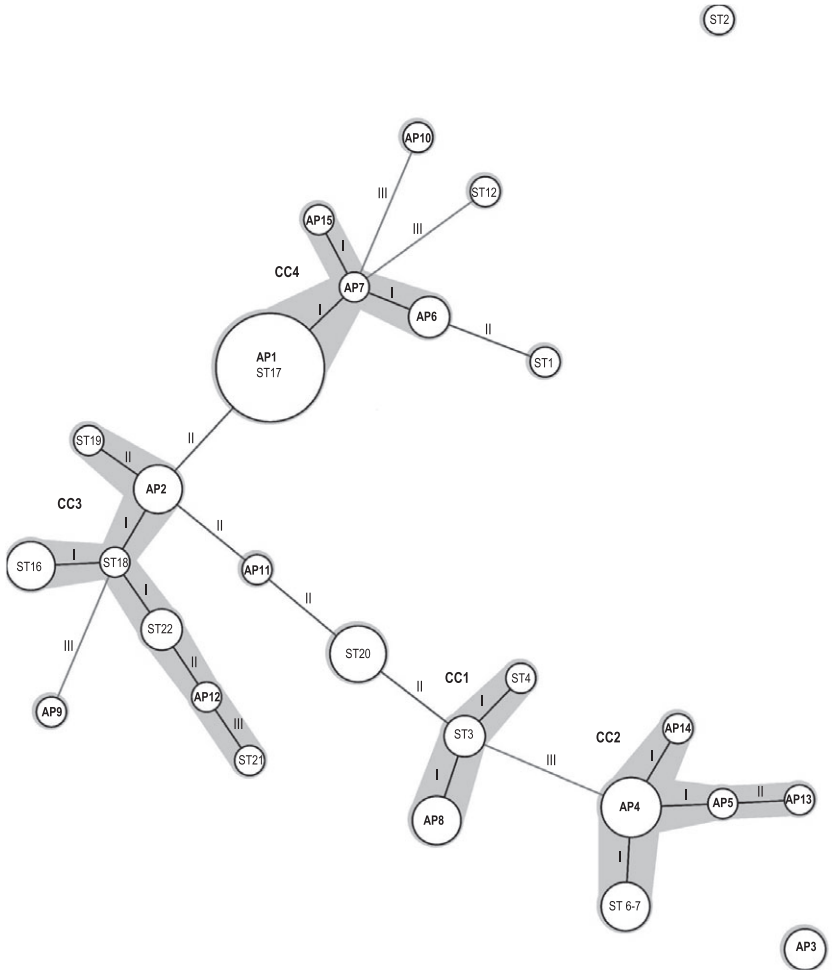

Figure 1 Minimum spanning tree of 32 *Xanthomonas campestris* pv. *campestris* (Xcc) strains analysed in this study (31 from Italy and one reference strain) and 30 Xcc strains from an international collection (Fargier *et al.*, 2011), based on multilocus sequence type analysis of seven loci using BioNumerics v. 7.1. Each circle indicates an allelic profile detected within Italian Xcc population (AP, in bold), or a sequence type delineated within an international collection (ST; Fargier *et al.*, 2011). The size of the circle is proportional to the number of strains. Grey zones surround different genotypes that belong to the same clonal complex. Black lines connecting pairs of APs/STs indicate that they varied at one locus (I), two loci (II), or three loci (III).

reported in the phylogenetic tree built using the neighbour-joining method.

Race determination

A set of eight differential *Brassica* lines was used for the race-typing of 31 strains of Xcc isolated in Italy. The response to inoculations varied depending on the relationship between the *Brassica* lines and Xcc strains according to the gene-for-gene model postulated by Vicente & Holub (2013). Reference strains (WHRI 3811, WHRI 1279A, WHRI 6181) showed the virulence pattern of races 1, 4 and 6 in accordance with their previous race designation (Vicente *et al.*, 2001; Fargier & Manceau, 2007). All Xcc strains isolated in Italy belonged to races 1, 4 and 6 (Table 1). Out of 31 strains of Xcc, 10 strains isolated from cauliflower, broccoli, kale, cabbage, kohlrabi and rutabaga were scored as race 1 (32.3%), 17 from cauliflower, kale, cabbage, Savoy cabbage, kohlrabi and seakale as race 4 (54.8%) and four from cauliflower, broccoli and kale as race 6 (12.9%; Table 3). Race 4 was identified in all the regions of Italy except Emilia Romagna, from which only one strain was analysed. Seven out eight Xcc strains isolated in Italy from kohlrabi (*Brassica oleracea* var. *gongyloides*) belonged to race 4 (Table 3). The races were distributed between the two main phylogenetic clusters I and IIA. All race 6 strains (four Italian and one reference), seven Italian race 4 strains and eight Italian race 1 strains were in cluster I, whereas cluster IIA contained 10 Italian race 4 strains and two Italian race 1 strains.

Discussion

Black rot of Brassicaceae caused by Xcc leads to severe yield losses in Italy, where recurrent outbreaks have been reported (Caponero & Iacobellis, 1994; Scortichini *et al.*, 1994; Catara *et al.*, 1999; Buonaurio *et al.*, 2003; Zaccardelli *et al.*, 2008).

A previous study based on analysis of DNA fingerprints obtained using primers based on DNA repetitive sequences, highlighted the high variability of the Italian Xcc population from central and southern Italy (Zaccardelli *et al.*, 2008). The 141 strains analysed by Zaccardelli *et al.* (2008) were used as a starting point to

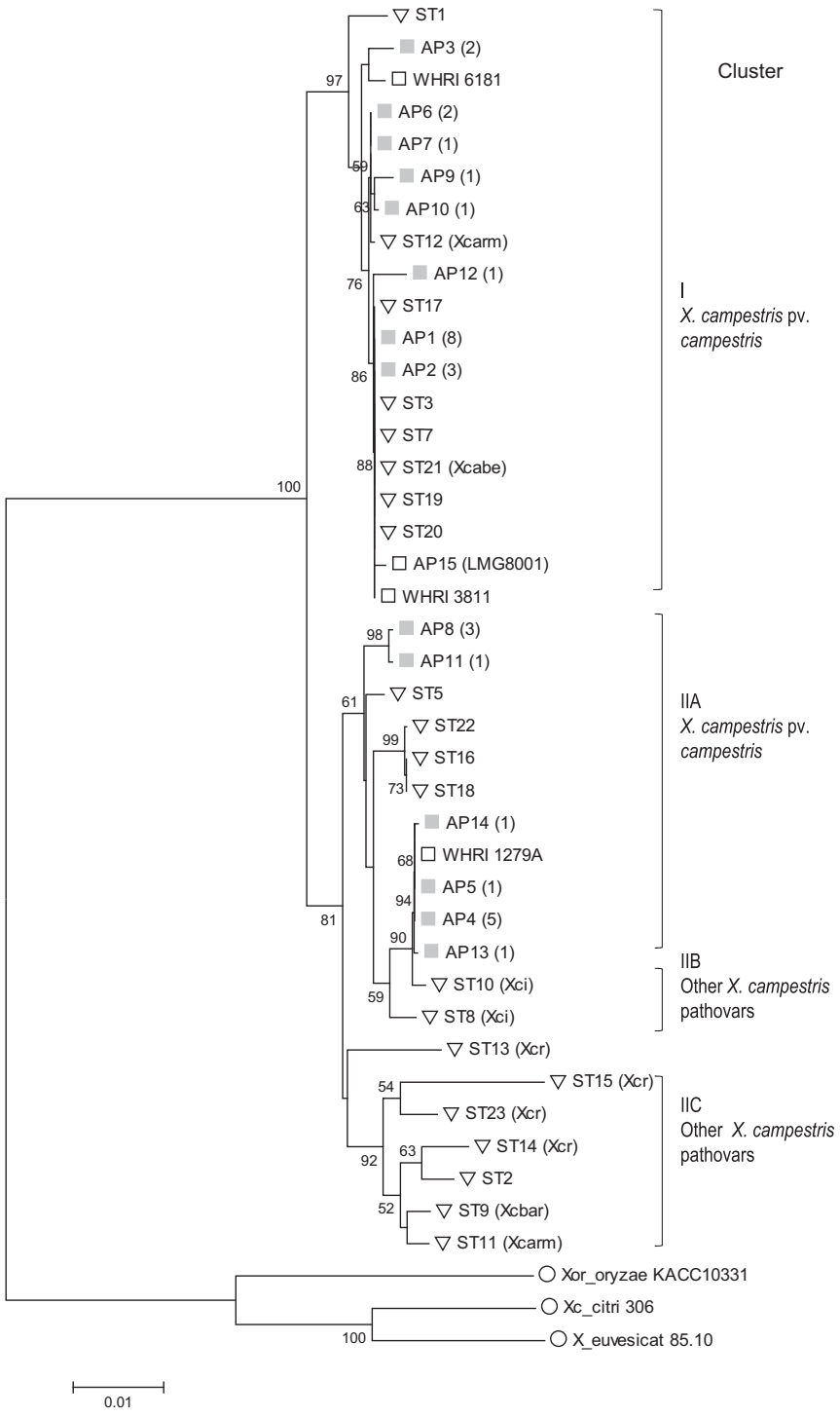

Figure 2 Phylogenetic tree of the concatenated nucleotide sequences of four genes, *dnaK*, *fuyA*, *gyrB* and *rpoD*, of *Xanthomonas campestris* pv. *campestris* (Xcc) strains analysed in this study and *X. campestris* pathovars of a worldwide origin from Fargier *et al.* (2011). The number of strains grouped in each allelic profile (AP) or sequence type (ST) is shown in parentheses. The evolutionary history was inferred using the neighbour-joining method and the distances were computed using the Jukes–Cantor method. Three out-groups were also included in the analysis. Bootstrap scores greater than 50 are displayed at each node. Grey square, Italian Xcc strains; white square, Xcc strains from other collections analysed in this study; white triangle, *X. campestris* pathovars from Fargier *et al.* (2011); white circle, out-group strains. Xcc, *Xanthomonas campestris* pv. *campestris*; Xcarm, *X. campestris* pv. *armoraciae*; Xcr, *X. campestris* pv. *raphani*; Xci, *X. campestris* pv. *incanae*; Xcabe, *X. campestris* pv. *aberrans*; Xo oryzae, *X. oryzae* pv. *oryzae;* Xc citri, *X. citri* subsp. *citri*; X euvesicat, *X. euvesicatoria*.

Table 3 Distribution of the three races identified within Italian *Xanthomonas campestris* pv. *campestris* strains according to the host of isolation.

Host	No. of strains	Race 1	Race 4	Race 6
Cauliflower	9	4	3	2
Kohlrabi	8	1	7	0
Cabbage	5	2	3	0
Kale	4	1	2	1
Broccoli	2	1	0	1
Savoy cabbage	1	0	1	0
Rutabaga	1	1	0	0
Seakale	1	0	1	0
Total	31	10	17	4
%		32.3	54.8	12.9

select strains representative of the genotypic diversity together with the origin of the strains in terms of host and region of isolation.

Here, strains were studied by an MLST approach based on seven genes (*atpD*, *dnaK*, *efp*, *fyuA*, *glnA*, *gyrB* and *rpoD*) and were race typed by inoculation on eight differential *Brassica* lines. All loci sequenced were polymorphic and, unlike those reported by Fargier *et al.* (2011), new alleles were identified in Italian Xcc strains at five out of seven loci (*dnaK*, *efp*, *fyuA*, *glnA* and *rpoD*), giving rise to 14 allelic profiles regardless of the geographical origin. Thus, based on the number of allelic profiles generated by these seven loci, bacterial populations observed in Italy were, in terms of diversity, as variable as those of a worldwide collection of 30 Xcc strains (Fargier *et al.*, 2011). AP1 was common to Xcc strains isolated in five out of seven Italian regions and was the most widespread amongst the Italian Xcc strains. This genotype was also the most common in the international collection of Xcc strains (ST17 in Fargier *et al.*, 2011) and was also found in 13 strains isolated in New York State (Lange *et al.*, 2016), thus indicating that it is has a worldwide distribution.

In addition to the three clonal complexes already described by Fargier *et al.* (2011), the AP1/ST17 strains form a new clonal complex (CC4) with other genotypes identified in strains isolated in Italy. Italian genotypes were present in all four clonal complexes. The founder genotype of CC2 and CC4 was also detected in one or more Italian strains. These clonal complexes included Italian genotypes linked by a single locus variation that in turn were linked to singletons by double locus variations. CC4 included the only genotype (i.e. AP1/ST17) shared both by Italian Xcc strains and those of a worldwide origin. This genotype was also identified in a strain isolated in 1967 (Fargier *et al.*, 2011). Most of the strains belonging to this genotype (i.e. AP1/ST17) were characterized as race 1. MLST suggests that genetically similar or related strains of genotype AP1/ST17 isolated in different continents over a long period of time are probably dispersed worldwide due to the brassica seed trade. The strains of the clonal complexes CC2 and CC4, on the other hand, could represent Xcc bacterial populations that have evolved locally. In CC2, a prevalent genotype AP4, which was not described by Fargier *et al.*, 2011, and three other minor genotypes were found. Genotypes encountered in other Italian strains in CC1 and CC3 or singletons could have evolved from isolated introductions. Different allelic profiles were detected in the different geographical areas except for the three strains isolated from Lazio that all showed the same allelic profile (AP8), unique for this region, although they were obtained from different *B. oleracea* varieties.

Some Xcc strains isolated from the same cabbage variety in several New York State counties in 2004 showed the same BOX-PCR profile and MLSA haplotype; however, this correlation was not found in all *X. campestris* strains (Lange *et al.*, 2016). In the present study, Xcc strains from different origins included in the same allelic profile, belonged to different M13 and BOX haplotypes (Zaccardelli *et al.*, 2008).

The phylogenetic reconstruction using the neighbour-joining algorithm with a concatenated dataset of the four loci (*dnaK*, *fuyA*, *gyrB* and *rpoD*) showed that Italian strains belonged to two genetic groups: cluster I that included only Xcc strains isolated in Italy and worldwide, or subcluster IIA, being separated from other pathovars of *X. campestris* that formed clearly distinct subclusters.

Recombination events evident in the split decomposition graph of the concatenated sequence of four loci confirmed previous findings (Fargier *et al.*, 2011). No correlation was found between clusters, host of isolation and geographical origin of the isolates, because the Xcc strains included in the two clusters were isolated from different Italian regions and plant hosts. Two genetic groups were also highlighted in an international collection of Xcc strains (Fargier *et al.*, 2011) and with an Xcc population from New York State (Lange *et al.*, 2016) in which Xcc strains were split into phylotypes that were not linked to origin and host of isolation.

Three physiological races, namely 1, 4 and 6, were identified within the Italian Xcc population based on the response to the inoculation of eight differential *Brassica* lines (Vicente *et al.*, 2001; Fargier & Manceau, 2007; Vicente & Holub, 2013). Several studies have shown that differences in relative distribution and frequencies of Xcc races were related to geographical area of isolation (Vicente *et al.*, 2001; Vicente, 2004; Jensen *et al.*, 2010; Lema *et al.*, 2012; Mulema *et al.*, 2012; Bila *et al.*, 2013; Rathaur *et al.*, 2015; Singh *et al.*, 2016).

Race 4 was predominant within the Italian Xcc strains (54.8%), followed by race 1 (32.3%). These two races were identified in six out seven Italian regions. Race 4 was also the main race reported in other countries such as northwestern Spain, Portugal, East Africa and Nepal (Vicente, 2004; Jensen *et al.*, 2010; Lema *et al.*, 2012; Mulema *et al.*, 2012). In contrast, in the UK and India, race 1 was predominant (Vicente & Holub, 2013; Rathaur *et al.*, 2015; Singh *et al.*, 2016). In southern Mozambique, all strains isolated from seed lots and

cabbage plants with symptoms and Portuguese tronchuda coles were identified as race 1 (Bila *et al.*, 2013). Only four Italian Xcc strains were characterized as race 6 (12.9%). In Portugal, Spain, Nepal and India, this race has also been found at low frequency (Vicente *et al.*, 2001; Jensen *et al.*, 2010; Lema *et al.*, 2012; Rathaur *et al.*, 2015; Singh *et al.*, 2016). Vicente *et al.* (2001) found no correlation between race and geographical origin of the isolates; however, some evidence for a relationship between race and host of origin was found. Race 1, 2, 3 and 4 isolates were all from different *B. oleracea* crops or cruciferous weeds growing in association with *B. oleracea* crops (Vicente *et al.*, 2001). In the present study, most Italian Xcc strains were isolated from *B. oleracea* and belonged to races 1 and 4; however, they were identified in all the regions analysed, thus indicating no correlation between race and geographic origin. Furthermore, Xcc strains belonging to races 1 and 4 were grouped independently in the two clusters, whereas Xcc strains belonging to race 6 were included only in cluster I. Jensen *et al.* (2010) analysed strains from Nepal and indicated that races 1, 4 and 6 were the most common in cabbage, whereas all strains from cauliflower belonged to race 4. In the present study, there was no relationship between *B. oleracea* varieties and Xcc races, except that most Xcc strains isolated from *B. oleracea* var. *gongylodes* in three regions belonged to race 4.

Xcc is a seedborne pathogen; however, it can survive in soil both as free cells and, for longer periods, associated with plant residues (Vicente & Holub, 2013). Infected seeds and transplants are the main way that Xcc is dispersed in the field. The current nationwide population study in Italy has shown that different Xcc populations coexist. Xcc strains sharing an allelic profile (AP1) with those from a worldwide origin have probably been introduced by infected traded seeds as also observed in Italy for other phytopathogenic bacteria (Bella *et al.*, 2012; Ialacci *et al.*, 2016). However, local endemic Xcc populations could have evolved in Italy as the result of an agricultural production system in which the farmers self-produce *Brassica* landrace seeds, thus perpetuating the same bacterial genotype within a field or home garden over several years.

Acknowledgements

This work was funded by Regione Siciliana 'Progetto Sementiero della Regione Siciliana' and supported by FIR 2014 (project code 1527F0) of the University of Catania 'Valorization of glucosinolates-myrosinase system to control pathologies of living organisms'. The authors declare no conflict of interest.

References

Bella P, Ialacci G, Licciardello G, La Rosa R, Catara V, 2012. Characterization of atypical *Clavibacter michiganensis* subsp. *michiganensis* populations in greenhouse tomatoes in Italy. *Journal of Plant Pathology* **94**, 635–42.

Bila J, Mortensen N, Andresen M, Vicente JG, Wulff EG, 2013. *Xanthomonas campestris* pv. *campestris* race 1 is the main causal agent of black rot of brassicas in southern Mozambique. *African Journal of Biotechnology* **12**, 602–10.

Branca F, 2008. Cauliflower and broccoli. In: Prohens J, Nuez F, eds. *Vegetables I*. New York, NY, USA: Springer, 147–82.

Buonaurio R, Santangeli F, Moretti C, Caglioti C, 2003. Occurrence of black rot caused by *Xanthomonas campestris* pv. *campestris* on ornamental kale in Italy. *Journal of Plant Pathology* **85**, 63.

CABI, 2017. *Xanthomonas campestris* pv. *campestris* (black rot). Invasive species compendium. Wallingford, UK: CAB International. [https://www.cabi.org/isc/datasheet/56919]. Accessed 17 August 2017.

Caponero A, Iacobellis NS, 1994. Foci of 'black rot' on cauliflower in Basilicata. *Informatore Agrario* **50**, 67–8.

Catara V, Branca F, Bella P, 1999. Epidemie di 'marciume nero' delle *Brassicaceae* in Sicilia. *Informatore Fitopatologico* **48**, 7–10.

Ciancaleoni S, Chiarenza GL, Raggi L, Branca F, Negri V, 2014. Diversity characterization of broccoli (*Brassica oleracea* L. var. *italica* Plenck) landraces for their on farm (*in situ*) safeguard and use in breeding programs. *Genetic Resources and Crop Evolution* **61**, 451–64.

Cruz J, Tenreiro R, Cruz L, 2017. Assessment of diversity of *Xanthomonas campestris* pathovars affecting cruciferous plants in Portugal and disclosure of two novel *X. campestris* pv. *campestris* races. *Journal of Plant Pathology* **99**, 403–14.

Dixon GR, 2007. *Vegetable Brassicas and Related Crucifers*. Crop Production Science in Horticulture Series No. 14. Wallingford, UK: CABI.

FAOSTAT, 2014. FAOSTAT. Crops. [http://www.fao.org/faostat/en/#data/QC]. Accessed 12 October 2017.

Fargier E, Manceau C, 2007. Pathogenicity assays restrict the species *Xanthomonas campestris* into three pathovars and reveal nine races within *X. campestris* pv. *campestris*. *Plant Pathology* **56**, 805–18.

Fargier E, Saux MFL, Manceau C, 2011. A multilocus sequence analysis of *Xanthomonas campestris* reveals a complex structure within crucifer-attacking pathovars of this species. *Systematic and Applied Microbiology* **34**, 156–65.

Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG, 2004. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *Journal of Bacteriology* **186**, 1518–30.

Huson DH, 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**, 68–73.

Huson DH, Bryant D, 2006. Application of phylogenetic networks in evolutionary studies. *Molecular Biology and Evolution* **23**, 254–67.

Ialacci G, Bella P, Licciadello G *et al.*, 2016. Clonal populations of *Clavibacter michiganensis* subsp. *michiganensis* are responsible for the outbreaks of bacterial canker in greenhouse tomatoes in Italy. *Plant Pathology* **65**, 484–95.

ISTAT, 2016. Agricoltura e zootecnia. Istituto nazionale di statistica. [http://agri.istat.it/sag_is_pdwout/jsp/Introduzione.jsp?id=21A|15A]. Accessed 12 October 2017.

Jensen BD, Vicente JG, Manandhar HK, Roberts SJ, 2010. Occurrence and diversity of *Xanthomonas campestris* pv. *campestris* in vegetable brassica fields in Nepal. *Plant Disease* **94**, 298–305.

Lange HW, Tancos MA, Carlson MO, Smart CD, 2016. Diversity of *Xanthomonas campestris* isolates from symptomatic crucifers in New York State. *Phythopathology* **106**, 113–22.

Lema M, Cartea ME, Sotelo T, Velasco P, Soengas P, 2012. Discrimination of *Xanthomonas campestris* pv. *campestris* races among strains from northwestern Spain by *Brassica* spp. genotypes and rep-PCR. *European Journal of Plant Pathology* **133**, 159–69.

Licciardello G, Bella P, Catara V, 2011. Quantitative detection of *Pseudomonas corrugata* and *P. mediterranea* in tomato plants by duplex real-time PCR. *Journal of Plant Pathology* **93**, 595–602.

Mulema JK, Vicente JG, Pink DAC *et al.*, 2012. Characterisation of isolates that cause black rot of crucifers in East Africa. *European Journal of Plant Pathology* **133**, 427–38.

Rathaur PS, Singh D, Raghuwanshi R, Yadava DK, 2015. Pathogenic and genetic characterization of *Xanthomonas campestris* pv. *campestris* races based on rep-PCR and multilocus sequence analysis. *Journal of Plant Pathology and Microbiology* **6**, 317.

Scortichini M, Rossi MP, Ruggini L, Cinti S, 1994. Recurrent infections of *Xanthomonas campestris* pv. *campestris* on Cruciferae in some areas of central-southern Italy. *Informatore Fitopatologico* **44**, 48–50.

Singh D, Rathaur PS, Vicente JG, 2016. Characterization, genetic diversity and distribution of *Xanthomonas campestris* pv. *campestris* races causing black rot disease in cruciferous crops of India. *Plant Pathology* **65**, 1411–8.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA 6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution* **30**, 2725–9.

Vicente JG, 2004. A podridao negra das Cruciferas. In: Lopes G, ed. *COTHN Centro Operativo e Tecnologico Hortofruticola Nacional*. Alcobaça, Portugal, Alcobaca: COTHN Centro Operativo e Tecnologico Hortofruticola, 102.

Vicente JG, Holub EB, 2013. *Xanthomonas campestris* pv. *campestris* (cause of black rot of crucifers) in the genomic era is still a worldwide threat to brassica crops. *Molecular Plant Pathology* **14**, 2–18.

Vicente JG, Conway J, Roberts SJ, Taylor JD, 2001. Identification and origin of *Xanthomonas campestris* pv. *campestris* races and related pathovars. *Phytopathology* **91**, 492–9.

Williams PH, 1980. Black rot: a continuing threat to world crucifers. *Plant Disease* **64**, 736–42.

Zaccardelli M, Campanile F, Spasiano A, Merighi M, 2007. Detection and identification of the crucifer pathogen, *Xanthomonas campestris* pv. *campestris*, by PCR amplification of the conserved Hrp/type III secretion system gene *hrcC*. *European Journal of Plant Pathology* **118**, 299–306.

Zaccardelli M, Campanile F, Moretti C, Buonaurio R, 2008. Characterization of Italian populations of *Xanthomonas campestris* pv. *campestris* using primers based on DNA repetitive sequences. *Journal of Plant Pathology* **90**, 375–81.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.

Figure S1. Split decomposition analysis based on concatenated sequences of four loci, *dnaK*, *fyuA*, *gyrB* and *rpoD* from the 15 allelic profiles detected within *Xanthomonas campestris* pv. *campestris* strains from Italy, reference strain (LMG8001) and 21 sequence types obtained from GenBank (Fargier *et al.*, 2011).

Table S1. Alleles with their GenBank accession numbers identified at seven loci by Fargier *et al.* (2011) and used to compare the sequences and define the allelic profiles of *Xanthomonas campestris* pv. *campestris* strains analysed in this study.

Table S2. Sequence types and general data for *Xanthomonas campestris* strains retrieved from Fargier *et al.* (2011), sequences of which were used to infer phylogenetic analysis of Italian *X. campestris* pv. *campestris* strains.

Table S3. GenBank accession numbers of the nucleotide sequences of alleles at seven loci, *atpD*, *dnaK*, *fyuA*, *efP*, *gyrB*, *glnA* and *rpoD*, of the *Xanthomonas campestris* pv. *campestris* strains analysed in this study.

Table S4. Clonal complexes identified within *Xanthonomas campestris* pv. *campestris* strains from an international collection (Fargier *et al.*, 2011) and from *X. campestris* pv. *campestris* strains used in this study.